Using Local Dictionaries

Structure of Local Dictionaries

Like the PDS4 master schema, local dictionaries will usually come in two parts: the XML Schema Definition language, or *.xsd*, file, which contains the structural class and attribute information; and the Schematron language, or *.sch*, file, which contains standard value lists and attribute relationship requirements. Together, these two files constitute a dictionary that defines a specific and unique *namespace*. When you wish to use classes from this namespace in your PDS4 label, you will need to reference both of these files.

Where to Find Dictionaries

The officially configured and released dictionaries are available from this website at PDS:

https://pds.nasa.gov/datastandards/dictionaries/

Development versions should be used with caution and are available elsewhere on the PDS4 site if desired.

Note: *In general, you should* ***not*** *be using mission dictionaries for missions other than the mission you are actually working on/for. If this presents a problem, contact your node consultant.*

Schema Files

To use any particular dictionary, you will need to download both the *.xsd* and the *.sch* files. If you've set up a local schema archive, you can copy these files into the appropriate place and continue.

Some software can also reference schema files across an available network connection, so it may not be necessary to have local copies of the files physically present on your system. You will need to know the specific URL, however, and you will also need set up your software to resolve references over a network connection.

Namespaces

A dictionary, that is, the combination of an *.xsd* file and its related *.sch* file, defines all the classes and attributes constituting a unique namespace. The formal name of that namespace will be a URI (Uniform Resource Identifier). For PDS4 namespaces, URIs are in the form of a URL. PDS maintains the dictionary collection at that URL, at least for now (there is no requirement that a URL-like URI be resolvable, but it does make life easier).

Finding the Dictionary Namespace URI

In order to reference the local dictionary in your labels, you will need to know the full, formal namespace URI.

To find the full namespace URI of a local dictionary, open the .xsd file and look at the contents of the `<xs:schema>` tag near the top of the file. Here, for example, is `<xs:schema>` tag from the version 1.1.0.0 Display Dictionary:

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  targetNamespace="http://pds.nasa.gov/pds4/disp/v1"
  xmlns:disp="http://pds.nasa.gov/pds4/disp/v1"
  xmlns:pds="http://pds.nasa.gov/pds4/pds/v1"
  elementFormDefault="qualified"
```

```
attributeFormDefault="unqualified"
version="1.1.0.0">
```

The URI of the namespace for any local dictionary is the value of the targetNamespace attribute in this tag. In the example above, we can see that the URI for the Display Dictionary namespace is "http://pds.nasa.gov/pds4/disp/v1".

The Standard Namespace Abbreviation

PDS also strongly prefers that you use standard namespace abbreviations for local dictionaries if you plan to reference namespaces by prefix (explained below). You can find the standard abbreviation in the <xs:schema> tag as well. Look for the xmlns: attribute that has the same value as the namespace URI for the dictionary, and the abbreviation will be the part following the colon. In the example above, the standard abbreviation for the Display Dictionary namespace is "disp".

Setting Up for Validation

If you have set up a local schema tree for referencing via an XML catalog file (see Understanding XML Catalog Files for details), or just so you can remember where they are, you will need to copy both the *.xsd* and *.sch* files into the appropriate place for your work environment.

The Tricky Part

Local dictionaries, once developed, tend to remain static. In particular, they are not necessarily updated every time the PDS4 master schema is updated. This is good in terms of keeping your label references current, but it does have a downside for validation. The local dictionaries will likely be referencing older versions of the PDS4 master schema. (Local dictionaries reference the master schema mainly for data type lists and unit classes, which are stable, so there is little reason to worry about updating schema references.)

But, if you want to be able to do full validation with your schema-aware editor, then you may have validation issues with the local dictionaries if you don't download the older versions of the master schema that they depend on. To see which version of the PDS4 master schema is being referenced, look for the <xs:import> statement near the top of the label. Here is the <xs:import> statement from the Display Dictionary:

```
<xs:import namespace="http://pds.nasa.gov/pds4/pds/v1"
  schemaLocation="http://pds.nasa.gov/pds4/pds/v1/PDS4_PDS_1101.xsd"/>
```

The version number of the master schema is part of this file name referenced in the schemaLocation attribute (here given as a URL reference). The PDS4 master schema version referenced here is "1.1.0.1". So to complete the validation chain you will need to download the *.xsd* and *.sch* files for this older version of the master schema, and install them in your local environment as well.

This is where a local schema library and XML Catalog file can make your validation life much easier. See Understanding XML Catalog Files for more information.

Referencing Local Dictionaries in PDS4 Labels

To some extent, how you reference the dictionaries depends on your work environment for creating and validating labels. Different validating editors will have different methods for finding and referencing the *.xsd* and *.sch* files, and these will not necessarily be the same, or consistent from editor to editor or even OS to OS. So following are some general hints and tips, but you should expect to have to either read some documentation or do some experimenting to find the optimal solution for yourself.

Schematron Reference

In order to validate against the contents of the *.sch* file, you will need to include a reference to it at the top of your label *.xml* file. The Schematron references will immediately follow the <?xml> line at the top of your label file. Here's an example from a label that references both the PDS4 master Schematron file, and the Display Dictionary Schematron file:

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="http://pds.nasa.gov/pds4/pds/v1/PDS4_PDS_1201.sch" schematypens="http://
purl.oclc.org/dsdl/schematron"?>
<?xml-model href="http://pds.nasa.gov/pds4/disp/v1/PDS4_DISP_1100.sch" schematypens="http://
purl.oclc.org/dsdl/schematron"?>
```

A few things to note:

- There is one xml-model line for each Schematron file to be referenced.
- The href reference in this case is to a URL based on the URI for the namespace. This will work if your software can reference the URL over the network, or if your software can use an XML Catalog file to turn the above reference into a local disk reference. If neither of these is possible, you will likely need to put a hard-coded reference to a disk file in here instead. The format for that may depend on your software as well.
- The schematypens ("schema type namespace") provides the URI of the namespace for the type of schema being referenced, in this case the ISO Schematron standard namespace. This is optional. As with many of these external reference techniques, some software may require this option be included, some may ignore it, and some may choke on it.
- xml-model is a processing instruction defined by the W3C group for referencing schemas. It has other attributes not shown here that may be considered vital to your software, or that you might see in other labels.

Because it is a defined standard, the xml-model processing instruction is the preferred form of reference for PDS4 archival labels, as opposed to editor-specific processing instructions that might do the same thing.

Namespace definition

To reference the contents of the *.xsd* file associated with the local dictionary, you will have to declare that you intend to use the namespace. You can also define an abbreviation for it and give its location in the same place. This can all be done in the root document tag - so, for example, if you are labelling an observation, it will be in the <Product_Observational> opening tag. Here's an example from a label that references both the master schema and the Display Dictionary schema:

```
<Product_Observational xmlns="http://pds.nasa.gov/pds4/pds/v1"
   xmlns:disp="http://pds.nasa.gov/pds4/disp/v1"
   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xsi:schemaLocation="http://pds.nasa.gov/pds4/pds/v1 http://pds.nasa.gov/pds4/pds/v1/
PDS4_PDS_1201.xsd
   http://pds.nasa.gov/pds4/disp/v1 http://pds.nasa.gov/pds4/disp/v1/PDS4_DISP_1100.xsd">
```

Lots of things going on here:

- The xmlns attribute is used to declare the full, formal namespace URI of a namespace (i.e., dictionary) we intend to use in this label in some way.
- The first xmlns attribute, without a colon and abbreviation following it, indicates that unless otherwise stated, all classes and attributes come from the *http://pds.nasa.gov/pds4/pds/v1* namespace - that is, the PDS4 master schema namespace.
- The xmlns:disp attribute assigns the standard abbreviation "disp" to all classes and attributes from the Display Dictionary namespace. So now if we want to use the *Display_Settings* class from that dictionary, we can use the "disp" prefix to reference it as <disp:Display_Settings>.
- The xmlns:xsi attribute invokes a standard namespace (probably built into your schema-aware editor so you don't need to supply a schema to define it) that makes the xsi:schemaLocation attribute available for use.
- The xsi:schemaLocation attribute contains a list of "*namespace_URI location*" pairs that give software a hint where they can find the namespace definition. In this case you see URLs that the editor in question was able to resolve. As with the <?xml-model> processing instruction, though, what sort of value will work for you depends on your environment. In some environments, the xsi:schemaLocation hint might not be needed at all.

Do *not* use the <xs:import> tag that you might see in the dictionary *.xsd* files to reference dictionary namespaces in your PDS4 labels. Importing only works inside an XSD schema file, and it is different from the referencing process we want here in the labels.

Using Classes from a Namespace

Once the references at the top of label file are squared away (you can see details on how to do this in Eclipse, for example, in our Using Eclipse for XML Editing pages), you can reference classes from the local dictionaries in the appropriate locations of your PDS4 label:

- Discipline dictionaries, like the Display Dictionary or Spectral Dictionary, can be referenced in the <*Discipline_Area*> of your label.
- Mission dictionaries should be referenced from the <*Mission_Area*> of your label.

Both of these areas are located at the bottom of the <*Observation_Area*> in observational products, and the <*Context_Area*> in other types of products.

Note: While it is not strictly necessary to declare all the namespaces you plan to reference in the document root tag, it is convenient to have them listed there all in one place for human readers, especially if there is schemaLocation information included.

When referencing classes and attributes from a local dictionary, you will need to indicate the namespace that contains their definition. There are two common ways to do this: By prefix and by changing the default namespace in a local context. Which method you use is largely a matter of personal preference or development style; both are equally valid.

Referencing by Prefix

If you've set your PDS4 label up as described in the previous examples, so that the default namespace is the PDS4 master schema namespace and there are abbreviations defined for the dictionary namespaces, then you can simply prefix the local dictionary abbreviation to classes and attributes from that dictionary. Using this method for the *Display_Settings* class would look like this:

```
<Discipline_Area>


  <disp:Display_Settings>
```

```
    <disp:Local_Internal_Reference>
      <disp:local_identifier_reference>Image1</disp:local_identifier_reference>
      <disp:local_reference_type>display_settings_to_array</disp:local_reference_type>
    </disp:Local_Internal_Reference>
    <disp:Display_Direction>
      <disp:horizontal_display_axis>Sample</disp:horizontal_display_axis>
      <disp:horizontal_display_direction>Left to Right</disp:horizontal_display_direction>
      <disp:vertical_display_axis>Line</disp:vertical_display_axis>
      <disp:vertical_display_direction>Top to Bottom</disp:vertical_display_direction>
    </disp:Display_Direction>
  </disp:Display_Settings>


  ...


</Discipline_Area>
```

Note that the prefix must be used on closing tags as well as opening tags, and on *all* attribute and class tags, not just the outermost.

Referencing by Changing the Default

Alternately, you can declare a different default namespace for all the contents of a single tag by using an xmlns attribute for that tag. The equivalent version for the *Display_Settings* example above would look like this:

```
<Discipline_Area>


  <Display_Settings xmlns="http://pds.nasa.gov/pds4/disp/v1">
    <Local_Internal_Reference>
      <local_identifier_reference>Image1</local_identifier_reference>
      <local_reference_type>display_settings_to_array</local_reference_type>
    </Local_Internal_Reference>
    <Display_Direction>
      <horizontal_display_axis>Sample</horizontal_display_axis>
      <horizontal_display_direction>Left to Right</horizontal_display_direction>
      <vertical_display_axis>Line</vertical_display_axis>
      <vertical_display_direction>Top to Bottom</vertical_display_direction>
    </Display_Direction>
  </Display_Settings>


  ...


</Discipline_Area>
```

After the *</Display_Settings>* closing tag, the default namespace goes back to being the PDS4 master schema namespace.